Price discovery in fragmented markets

de Jong, F.C.J.M.; Schotman, P.C.

Published in:
Journal of Financial Econometrics

DOI:
10.1093/jjfinec/nbp015

Publication date:
2010

Journal of Financial Econometrics, 2010, Vol. 8, No. 1, 1–28

Price Discovery in Fragmented Markets

Frank de Jong
Tilburg University

Peter C. Schotman
Maastricht University

ABSTRACT

This paper proposes a structural time-series model for the intraday price dynamics on fragmented financial markets. We generalize the structural model of Hasbrouck (1993) to a multivariate setting. We discuss identification issues and propose a new measure for the contribution of each market to price discovery related to the Hasbrouck (1995) information shares. We apply the model to two sets of Nasdaq dealer quotes. (*JEL*: C32, F31)

KEYWORDS: High-frequency data, microstructure, structural time-series models

The markets in many financial assets are fragmented. To give a few examples, NYSE-listed U.S. stocks are often also traded on regional exchanges; many European stocks are cross-listed on the NYSE or Nasdaq; on Nasdaq itself and in the foreign exchange and bond markets, there are multiple dealers and the markets for the trading between dealers and their clients is quite separated from the interdealer market.

Price discovery models aim to describe the dynamic interactions between the quotes or transaction prices from two or more markets, or from two or more dealers of the same asset. Based on these dynamics, the relative contribution of each market or dealer to the price discovery process can be assessed.

We would like to thank Bart Frijns, Bruce Lehmann, and seminar participants at University of Amsterdam, Bilgi University (Istanbul), Stockholm School of Economics, University of Uppsala, and Sveriges Riksbanken for helpful comments on an earlier draft. As always, all errors are our own. Address correspondence to Peter Schotman, Limburg Institute of Financial Economics (LIFE), Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands, or e-mail: P.Schotman@MaastrichtUniversity.NL.

doi: 10.1093/jjfinec/nbp015
Advance Access publication September 10, 2009

The most natural model for prices p_{it} on market i (or quotes by dealer i) is that they equal the fundamental value of the asset, p_t^*, plus a transitory term:[1]

$$p_{it} = p_t^* + u_{it}. \tag{1}$$

In the Madhavan (2000) survey this model forms the basis to analyze trading frictions, asymmetric information, and inventory control. Equation (1) is in the form of an unobserved components model, or a structural time-series model in the econometrics terminology of Harvey (1989). Prices are observed, but the efficient price (or fundamental value) p_t^* is not. The fundamental value is assumed to be a random walk, whereas the market (dealer)-dependent transitory term u_{it} is stationary and typically close to white noise. The price changes, Δp_{it}, therefore have a very typical serial correlation pattern: a strong and negative first-order autocorrelation, and small and often negligible higher-order autocorrelations.

Despite its intuitive appeal, the unobserved components model is rarely used in empirical work, neither for estimation nor for the definition of measures of price discovery. Starting with Hasbrouck's (1995) pioneering work, the standard time-series model proposed by Hasbrouck (1995) is the Vector AutoRegression introduced by Sims (1980) in macroeconomics. Since all price series share the same long-term (random walk) component, the VAR is subject to cointegration restrictions and estimated as a vector error correction model (VECM). The central quantity of interest is the *information share*, which measures the relative importance of each market in the price discovery process. Hasbrouck (1995) defines the information share as the fraction of the variance of the random walk component that can be attributed to a particular market (or dealer). The VECM and information share methodology have been applied in many empirical studies.[2]

In this paper, we revisit the unobserved components microstructure model of Hasbrouck (1993) in Equation (1) and extend it to a multiple markets setting. The information flow is modeled through the simultaneous and lagged covariances between the "noise" terms in (1) and the innovations in the fundamental value. Within this model, we introduce a new measure of the contribution to price discovery. Unlike the traditional information share, which is defined within a reduced-form time-series model, the new measure is defined directly within the unobserved components model. Apart from its intuitive appeal as a model for financial market data, working directly within the unobserved components model has several other advantages over the VECM approach in settings with many markets or many dealers.

[1] See, for example, Hasbrouck (1993), Zhou (1996), Lehmann (2002), and Bandi and Russell (2006) for a variety of applications.

[2] For example, Hasbrouck (1995) and Harris et al. (2002) for U.S. equities traded on the NYSE and regional exchanges; Taylor (2008) for the various futures contracts on the S&P500; Hupperets and Menkveld (2001) for European equities cross-listed in the U.S.; Upper and Werner (2002) for the relation between the cash and futures market in German government bonds; De Jong, Mahieu, and Schotman (1998); Covrig and Melvin (2002) for the foreign exchange market; Mizrach and Neely (2008) for the U.S. Treasury market; and Dittmar and Yuan (2008) for the relation between corporate and sovereign bonds in emerging markets.

First, the particular pattern of autocorrelations in prices (or quotes) is difficult to describe with low-order autoregressive models. Autoregressions often require long lags to capture a strong negative first-order autocorrelation and a second autocorrelation that is almost zero. The reduced form of the unobserved components model is a vector-moving average process that can fit the data at least as well as the VECM. The VECM also suffers from lack of parsimony in the error correction part. In a model with N dealers, the cointegration restrictions lead to $N-1$ different error correction terms in each of the N equations. The parsimony of the unobserved components model has advantages both for the statistical inference as well as the definition of information shares.

Related to this is a potential problem with the data. Although microstructure time series have many observations, we do not always have that many observations for *all* markets (dealers). For example, the NYSE is much more active than its regional satellite markets; foreign exchange dealers are often at a few large banks; most Nasdaq quotes are issued by a handful of dealers and the electronic communication networks (ECN). In these circumstances the time series for a multivariate model of dynamic interactions is sampled at the pace of the slowest market (Harris et al., 2002) or with relatively long, fixed calendar intervals. This problem is particularly serious for large-dimensional systems, that is, a setting with multiple markets. When the number of dealers increases, the number of simultaneously available observations generally decreases, but the number of parameters in a VAR increases quadratically with the number of time series. In the unobserved components model it is also straightforward to deal with differences in observation period across markets, caused by holidays, missing data, etc.[3]

Finally, the VECM model has problems in the construction of information shares. These are not uniquely defined, but depend on the allocation of the covariance terms in the error covariance matrix. Hasbrouck (1995) suggests to report upper and lower bounds, obtained by different ordering of the markets. For a two-variable system these bounds are sometimes fairly narrow, but there are also applications (for example, Covrig and Melvin, 2002) where the bounds are very wide. In a high-dimensional system, the number of off-diagonal elements in the covariance matrix increases quadratically in N, and will eventually dominate the variance decomposition, so that it is difficult to obtain meaningful estimates of the information shares. Our proposed information share measure does not depend on an arbitrary way to split the correlation of the reduced-form error term over the markets, and will therefore remain meaningful in high-dimensional settings.

The unobserved components model is appealing in these situations, but has a drawback of its own. Since Equation (1) contains the efficient price as a latent variable, there is an inherent identification problem.[4] In the multivariate unobserved components model, which is of interest for price discovery in fragmented

[3]Estimation methods based on Kalman filters are especially appropriate here. See, for example, Menkveld, Koopman, and Lucas (2007).

[4]For the univariate version of the model, this identification problem is discussed in depth in Hasbrouck (1993).

markets, the identification problem turns out to be less severe. Full identification, and hence a unique value for the information shares, is achieved under plausible assumptions regarding the idiosyncratic term u_{it}.

The structure of this paper is as follows. First, we provide a theoretical investigation of the properties of the structural price discovery model and discuss the various identification rules. Next, we present our alternative measure for the contribution to price discovery. We then extend the structural model to higher-order dynamics and compare the implications of this model with the usual VECM approach. We examine the economic meaning of information shares within a stylized theoretical microstructure model. We finally present two empirical applications.

1 A STRUCTURAL TIME-SERIES MODEL

This section explores a structural time-series model for market microstructure and price discovery in fragmented markets. The model generalizes the univariate model of Hasbrouck (1993) to a multiple market setting. This section first reviews the results for a univariate pure random walk plus noise model. Then the model is extended to a multivariate random walk plus noise. In a later section, higher-order dynamics are introduced.

1.1 Univariate Model

Hasbrouck (1993) considers the univariate structural model for p_t, the logarithm of the price of a security:

$$\begin{aligned} p_t &= p_t^* + u_t, & & \\ p_t^* &= p_{t-1}^* + r_t, & \mathrm{Var}(r_t) &= \sigma^2, \\ u_t &= \alpha r_t + e_t, & \mathrm{Var}(e_t) &= \omega^2, \end{aligned} \tag{2}$$

where p_t^* is the unobserved efficient price (random walk) and u_t a transitory component. The shocks e_t and r_t are uncorrelated. The coefficient α determines the covariance between transitory and permanent shocks: $\mathrm{Cov}(u_t, r_t) = \alpha\sigma^2$.

We can write the price changes (returns) in this model as

$$\Delta p_t = r_t + \Delta u_t = (1+\alpha) r_t - \alpha r_{t-1} + \Delta e_t. \tag{3}$$

The auto-covariances of returns implied by this model are therefore

$$\gamma_0 = \mathrm{E}\left[\Delta p_t^2\right] = \sigma^2((1+\alpha)^2 + \alpha^2) + 2\omega^2, \tag{4a}$$

$$\gamma_1 = \mathrm{E}[\Delta p_t \Delta p_{t-1}] = -\sigma^2 \alpha (1+\alpha) - \omega^2. \tag{4b}$$

All higher-order covariances are zero, and therefore the reduced form of the structural model is a first-order Moving Average process in the price changes.

From the moment equations, the parameter σ^2 is uniquely identified as

$$\sigma^2 = \gamma_0 + 2\gamma_1. \tag{5}$$

The parameters α and ω^2 cannot be identified separately. Hence, some identifying restriction is necessary. We first define a range of admissible values for α. From the moment conditions we obtain

$$\begin{aligned}\omega^2 &= -\gamma_1 - \alpha(1+\alpha)\sigma^2 \\ &= -\gamma_0(\rho_1 + \alpha(1+\alpha)(1+2\rho_1)),\end{aligned} \tag{6}$$

where $\rho_1 = \gamma_1/\gamma_0$ is the first-order autocorrelation. For microstructure data, the first-order autocorrelation is typically negative, but bigger than $-\frac{1}{2}$. For the interpretation of the model, ω^2 must remain non-negative. This provides a bound on the admissible values of α. Equation (6) implies the inequality

$$-\sqrt{1-2\rho_1} \leq (2\alpha+1)\sqrt{1+2\rho_1} \leq \sqrt{1-2\rho_1}. \tag{7}$$

These intervals typically contain both positive and negative values for α. Boundary cases are $\rho_1 \to -\frac{1}{2}$, in which case α is not restricted at all, and $\rho_1 = 0$, in which case $-1 \leq \alpha \leq 0$. For a typical first-order autocorrelation $\rho_1 = -0.3$, we find the interval $-1\frac{1}{2} \leq \alpha \leq \frac{1}{2}$.

Two identifying restrictions are popular in the literature: the Beveridge–Nelson (BN) normalization ($\omega^2 = 0$) and the Watson normalization ($\alpha = 0$). The BN normalization is always admissible. For admissibility of the Watson normalization ($\alpha = 0$), we need a negative first-order autocorrelation.[5] Hasbrouck (1993) shows that the choice of normalization for α may have an important effect on the variance of the idiosyncratic term ($\text{Var}(u_t)$) in empirical applications. In the UC model, we can write the variance of the idiosyncratic term, using (2) and (6), as

$$\text{Var}(u_t) = \alpha^2\sigma^2 + \omega^2 = -\gamma_1 - \alpha\sigma^2. \tag{8}$$

The noise variance attains a lower bound when α is at its maximum value, which corresponds to the BN normalization.

This completes the summary of Hasbrouck's (1993) model. We now turn to a multivariate generalization of his model.

1.2 Multivariate Model

Let p_t now be a vector of N prices for the same asset from different markets. The multivariate model reads,

$$\begin{aligned} p_t &= \iota p_t^* + u_t, & & \\ p_t^* &= p_{t-1}^* + r_t, & \text{Var}(r_t) &= \sigma^2, \\ u_t &= \alpha r_t + e_t, & \text{Var}(e_t) &= \Omega, \end{aligned} \tag{9}$$

[5] Morley, Nelson, and Zivot (2003) study the identification of α in a model with positive first-order autocorrelation, which is typical for macroeconomic data. In that case, the range of admissible α may not contain zero, and the Watson restriction is not feasible. But since the first-order autocorrelation for microstructure return data is almost always negative, the Watson restriction is typically feasible for microstructure data.

where α is an N-vector, ι is a vector of ones, and Ω a $(N \times N)$ matrix. Again, $\mathrm{Cov}(u_t, r_t) = \alpha\sigma^2$. As in the univariate model, the innovations in the efficient price and the transitory term may be correlated. By construction, all price series share the same random-walk component and are therefore cointegrated. The price changes (returns) in this model are written as

$$\Delta p_t = \iota r_t + \Delta u_t = (\iota + \alpha) r_t - \alpha r_{t-1} + \Delta e_t, \tag{10}$$

which lead to the moment conditions

$$\Gamma_0 = \mathrm{E}[\Delta p_t \Delta p_t'] = \sigma^2((\iota + \alpha)(\iota + \alpha)' + \alpha\alpha') + 2\Omega, \tag{11a}$$

$$\Gamma_1 = \mathrm{E}[\Delta p_t \Delta p_{t-1}'] = -\sigma^2 \alpha (\iota + \alpha)' - \Omega. \tag{11b}$$

All parameters in this model are (over-)identified, except the vector α, which is only identified up to a translation along the unit vector. First, the sum of lead, current, and lag covariances,

$$\Gamma_1' + \Gamma_0 + \Gamma_1 = \sigma^2 \iota\iota', \tag{12}$$

overidentifies the variance of the efficient price innovation. Next, consider the difference between lead and lag cross-covariances:

$$\Gamma_1 - \Gamma_1' = \sigma^2(\iota\alpha' - \alpha\iota'). \tag{13}$$

From this, α can be identified up to a translation along ι. Finally, given the values for σ^2 and α, the noise covariance matrix Ω can be identified either from Equation (11a) or from the sum of the lead and lag covariances:

$$\Gamma_1 + \Gamma_1' = -\sigma^2(\alpha\iota' + \iota\alpha' + 2\alpha\alpha') - 2\Omega. \tag{14}$$

The entire set of equivalent solutions is characterized by

$$\alpha = \tilde{\alpha} - w\iota, \tag{15a}$$

$$\Omega = \tilde{\Omega} + w\sigma^2((1 - w)\iota\iota' + \iota\tilde{\alpha}' + \tilde{\alpha}\iota'), \tag{15b}$$

where w is an arbitrary scalar and $\tilde{\alpha}$ and $\tilde{\Omega}$ constitute an initial admissible solution. Since Ω is a covariance matrix, it must be positive semidefinite. Therefore, not all values for w are admissible, analogous to the univariate case. The range of alternative equivalent combinations of α and Ω in the multivariate model is smaller than in the univariate model. For each price series the univariate restrictions must hold for the diagonal element ω_{ii} and they must hold jointly. In addition, the positive definiteness for Ω is stronger than just positive diagonal elements.

Analogous to the univariate model, the BN representation provides an admissible solution $(\tilde{\alpha}, \tilde{\Omega})$. The BN representation is obtained from the reduced form. The reduced form of the multivariate random walk plus noise model is the first-order vector-moving average (VMA) process,

$$\Delta p_t = \epsilon_t - C\epsilon_{t-1}, \qquad \mathrm{Var}(\epsilon_t) = \Sigma, \tag{16}$$

where cointegration requires that

$$C = I - \iota\theta' \tag{17}$$

for some vector θ. The BN representation of the reduced form is

$$\begin{aligned} p_t &= \iota\tilde{p}_t + (I - \iota\theta')\epsilon_t, \\ \tilde{p}_t &= \tilde{p}_{t-1} + \theta'\epsilon_t. \end{aligned} \tag{18}$$

Under the BN restriction, the innovations in the permanent component are equal to an exact linear combination of the VMA innovations: $r_t = \theta'\epsilon_t$. Since the variance of the random-walk component is uniquely identified, we have

$$\sigma^2 = \theta'\Sigma\theta. \tag{19}$$

To relate the other parameters in the UC to the reduced-form parameters, we write

$$\mathrm{Cov}(\Delta p_t, r_t) = \Sigma\theta = \sigma^2(\iota + \alpha), \tag{20}$$

where the last equality follows from (10). This gives a particular choice for α that we shall call the BN value,

$$\tilde{\alpha} = \Sigma\theta/\sigma^2 - \iota. \tag{21}$$

For the BN normalization, the covariance matrix of e_t is semidefinite:

$$\tilde{\Omega} = \Sigma - \frac{\Sigma\theta\theta'\Sigma}{\theta'\Sigma\theta}. \tag{22}$$

All other normalizations of α and Ω are obtained from (15a) and (15b). In the Appendix we show that for $0 < \iota'\theta < 2$, only positive values for w are allowed. In that case the BN value of α is the maximal value, as it is in the univariate case.

For a generalization of the Watson restriction, we could assume that there is one market whose idiosyncratic term is uncorrelated with the efficient price, that is, by setting one element $\alpha_i = 0$. The interpretation of the Watson restriction is that one market is designated as the central market. In some applications there is a natural choice for the central market. For example, when studying the relation between the NYSE and regional markets in the U.S., the NYSE would be the central market. As another example, in an application with cross-listed stocks, the home market is the candidate central market. Setting some arbitrary $\alpha_i = 0$ could easily be inadmissable because it will violate the condition that Ω must be positive semidefinite. Admissability must be checked on a case by case basis and will restrict the potential normalizations of α. Imposing the Watson restriction $\alpha_i = 0$ on *every* market leads to $N-1$ over-identifying restrictions, which may be violated by the data.

In many applications, microstructure theory does not suggest a Watson-type normalization. An alternative identifying assumption is that Ω is diagonal. Under that assumption the deviations from the efficient price, $p_{it} - p_t^*$, will only be correlated across markets because of their joint dependence on the innovation in the efficient price r_t. Diagonality of Ω, of course, does not help identification in the univariate model. The bivariate case ($N = 2$) is special, since the off-diagonal element ω_{21} can be set to zero by a suitable choice of w in (15b) without imposing any further over-identifying conditions. When $N > 2$, assuming Ω is diagonal does put testable over-identifying restrictions on the data.

Diagonality of Ω is very different from diagonality of the reduced-form covariance matrix Σ. The latter is violated in any empirical application. With microstructure data, the typical covariances among price innovations are positive. In the UC these positive covariances are modeled by their common dependence on the efficient price using the coefficients $\beta = \iota + \alpha$. In the next section we analyze how the assumption facilitates the interpretation of information shares.

2 INFORMATION SHARES

Information measures of price discovery summarize the relation between the change in the efficient price and actual price changes. The most common measure is due to Hasbrouck (1995), who defines information shares within a reduced-form model. In the simplest case with only first-order dynamics, the VMA(1) model (16) from the previous section can be written in the permanent-transitory decomposition form (18), with $r_t = \theta' \epsilon_t$. Hasbrouck (1995) proposes the variance decomposition

$$\sigma^2 = \mathrm{Var}(r_t) = \theta' \Sigma \theta = \sum_{i=1}^{N} \sum_{j=1}^{N} \theta_i \theta_j \sigma_{ij} \tag{23}$$

to define information shares for each dealer. If the shocks ϵ_{it} would be mutually uncorrelated, the information shares

$$k_i = \frac{\theta_i^2 \sigma_{ii}}{\sigma^2} \tag{24}$$

would measure the part of the variance of the innovation to the efficient price that is due to the information in dealer i's quotes. When the covariances σ_{ij} are not equal to zero, it is not clear how much of the covariance $\theta_i \theta_j \sigma_{ij}$ should be attributed to dealers i and j. In empirical work the covariance terms are often large. For large N the covariance terms could even dominate the contributions of the diagonal elements. By varying the order of the variables in p_t in alternative Cholesky decompositions of Σ, it is possible to obtain an upper and a lower bound.

In this section we suggest a modification of this definition, which allocates the covariance terms in a particular way. Instead of the reduced-form definition, we define the information shares directly within the structural unobserved components model. From (9), price innovations in the UC model are given by

$$v_t = \iota r_t + u_t = (\iota + \alpha) r_t + e_t = \beta r_t + e_t, \tag{25}$$

and have covariance matrix

$$\mathrm{E}[v_t v_t'] = \Upsilon = \sigma^2 \beta\beta' + \Omega. \tag{26}$$

As in Hasbrouck (1995), we consider the relation between the innovation in the efficient price and the shocks to individual prices,

$$r_t = \gamma' v_t + \eta_t, \tag{27}$$

where η_t is the part of the innovation in the efficient price that is unrelated to innovations in observed prices. In the UC, η_t will generally have a positive variance, while in the reduced-form VMA, by construction, $\eta_t = 0$. The regression coefficients γ follow as

$$\gamma = \Upsilon^{-1} \beta \sigma^2. \tag{28}$$

Next, analogous to the Hasbrouck (1995) definition, consider a variance decomposition of r_t,

$$\mathrm{Var}(r_t) \equiv \sigma^2 = \gamma' \Upsilon \gamma + \sigma_\eta^2. \tag{29}$$

Because σ_η^2 is positive, not all variance can be attributed to innovations in observed prices. The total fraction of the variance in the fundamental price innovation r_t explained by the vector of observed price innovations is

$$R^2 = 1 - \sigma_\eta^2/\sigma^2 = \gamma' \Upsilon \gamma / \sigma^2 = \gamma' \beta = \sum_{j=1}^{N} \gamma_j \beta_j. \tag{30}$$

As information shares we propose

$$\mathrm{IS}_j = \gamma_j \beta_j. \tag{31}$$

For an interpretation of this definition, recall that β is the regression coefficient of the price innovations v_t on the efficient price r_t, while γ is the coefficient in the reverse regression of r_t on v_t. The product of the elements of these vectors can be interpreted as a partial R^2, indicating how much of the variance of r_t is explained by each element of v_t. These partial R^2s do not add up to one, because in the UC model some of the variation in the efficient price is uncorrelated with the observed price innovations.

The information shares are not invariant with respect to the normalization of α and Ω. Different choices for w will lead to different information shares. Without a credible choice of w the definition still contains some arbitrary allocation of covariances. As a plausible identification we consider the assumption that Ω is diagonal. In that case the only source of covariance between elements of v_t is through the common factor r_t. With Ω diagonal, we can express the information shares as in the following theorem (see the Appendix for proof).

Theorem 1. *Let information shares be defined by $IS_j = \beta_j \gamma_j$. Assume Ω diagonal with positive diagonal elements ω_j^2. Then*

$$IS_j = \frac{\beta_j^2/\omega_j^2}{1/\sigma^2 + \sum_i \beta_i^2/\omega_i^2}. \tag{32}$$

Information shares therefore depend on the ratio β_j/ω_j. The less the noise in market j, the higher the information share. Similarly, the stronger the covariance between prices in market j and the efficient price, the higher the information share.

Recall that $\beta_j = 1 + \alpha_j$. When a diagonal Ω is close to the Watson restriction with some central market having $\alpha_i = 0$, we expect that less informative satellite markets have $\alpha_j < 0$, or have a high ω_j. In other words, informationally less efficient markets will be characterized by slow or noisy price adjustment.

To see the relation between this definition of the information share and Hasbrouck's, consider first the BN normalization. From (21) it follows that $\tilde{\beta} = \Sigma\theta/\sigma^2$. This is also the highest possible value for β, because the BN normalization gives the highest possible value for α. By substituting the value of $\tilde{\Omega}$ from (22), we find that

$$\tilde{\Upsilon} = \sigma^2 \tilde{\beta}\tilde{\beta}' + \tilde{\Omega} = \Sigma. \tag{33}$$

Likewise, $\tilde{\gamma} = \tilde{\Upsilon}^{-1}\tilde{\beta} = \theta$. Hence, under the BN identification rule the information shares are

$$\tilde{\mathrm{IS}}_j = \tilde{\gamma}_j \tilde{\beta}_j = \frac{\sum_{i=1}^{N} \sigma_{ij}\theta_i\theta_j}{\sigma^2}. \tag{34}$$

By construction, these information shares add up to one. This is not surprising, since the variance of the residual in (27), σ_η^2, is zero in this case. These information shares are identical to Hasbrouck's (1995) definition if Σ is diagonal. In the generic case where Σ is not diagonal, this information share distributes the covariances between markets in a particular way.

3 EXAMPLE

Hasbrouck (2002) considers a number of stylized examples to evaluate the economic plausibility of alternative statistical price discovery measures. One of his examples concerns a simple version of the Glosten and Harris (1988) model. There are two markets, but all information is revealed in the first market. Prices are given by

$$\begin{aligned} p_{1t} &= p_t^* + q_{1t}, \\ p_{2t} &= p_{t-1}^* + q_{2t}, \\ p_t^* &= p_{t-1}^* + q_{1t}, \end{aligned} \tag{35}$$

where q_{1t} and q_{2t} are uncorrelated shocks with unit variance. To formulate this model in our notation, let $r_t = q_{1t}$, write

$$p_{2t} = p_t^* + p_{t-1}^* - p_t^* + q_{2t} = p_t^* - r_t + q_{2t},$$

let $e_{2t} = q_{2t}$, and set $e_{1t} = 0$. With this notation the dealer behavior can be written in the form of (9) as

$$\begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} r_t + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}, \tag{36}$$

where the vector $(e_{1t}\ e_{2t})$ has covariance matrix

$$\Omega = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{37}$$

From (36) it is immediate that $\sigma^2 = 1$ and $\alpha = (1 \quad -1)$. This is consistent with the moment conditions (13):

$$\alpha_1 - \alpha_2 = \mathrm{E}[\Delta p_{2t} \Delta p_{1,t-1}] - \mathrm{E}[\Delta p_{1t} \Delta p_{2,t-1}] = 2. \tag{38}$$

Given α, β follows as $(2 \quad 0)'$. The matrix Ω is diagonal in line with what we also think is the most plausible identification. For the parameter γ, we first compute the covariance matrix Υ from (26):

$$\Upsilon = \sigma^2 \beta\beta' + \Omega = \begin{pmatrix} 2 \\ 0 \end{pmatrix} (2 \quad 0) + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}. \tag{39}$$

Therefore, using (28),

$$\gamma = \Upsilon^{-1} \beta \sigma^2 = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}, \tag{40}$$

and the information shares follow from (31):

$$\begin{pmatrix} \mathrm{IS}_1 \\ \mathrm{IS}_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \odot \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{41}$$

exactly as intended by the example. Market 1 contains all the information and the information share IS_1 reflects this.

This should be a good point to leave the example, were it not that α is not uniquely identified from the data. Observationally equivalent representations arise by translating α along the unit vector, and doing a compensating transformation on Ω. The set of equivalent models in this example is (see (15a))

$$\begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \begin{pmatrix} 1 - w \\ -1 - w \end{pmatrix} r_t + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}, \tag{42}$$

Table 1 Observationally equivalent structural models.

w	α	Ω		β	γ	IS
0	1	0	0	2	$\frac{1}{2}$	1
	−1	0	1	0	0	0
0.25	0.75	0.688	0.188	1.75	0.456	0.799
	−1.25	0.188	0.813	−0.25	−0.155	0.039
3/5	0.40	1.440	0.240	1.40	0.200	0.280
	−1.60	0.240	0.040	−0.60	−1.200	0.720

The table reports alternative observationally equivalent parameter configurations of the model:

$$\begin{aligned} p_t &= p_t^* + u_t, \\ u_t &= \alpha r_t + e_t, \\ p_t^* &= p_{t-1}^* + r_t, \end{aligned}$$

related to the stylized example in Section 3.

with covariance matrix (see (15b))

$$\mathrm{E}[e_t e_t'] = \Omega = \begin{pmatrix} 3w - w^2 & w - w^2 \\ w - w^2 & 1 - w - w^2 \end{pmatrix}. \tag{43}$$

Hasbrouck's structural representation obtains for $w = 0$. Alternative representations are admissible if Ω is positive semidefinite. From the determinant of Ω and its diagonal elements, it follows that this restricts w to

$$0 \leq w \leq 3/5. \tag{44}$$

The minimal value $w = 0$ generates the BN representation.

Table 1 reports the implications of three representations corresponding to three different values of w. Results for the Hasbrouck identification ($w = 0$) have been discussed before. For the other observationally equivalent models, the noise covariance matrix Ω is not diagonal.[6] The relation between w and the information shares is far from linear.[7] For most values of w, like $w = 0.25$ in Table 1, market 1 remains dominant. The identification problem can, however, lead to a completely different interpretation as shown in the last rows ($w = 3/5$). In this case the market-specific shocks are not idiosyncratic at all, but perfectly correlated. Under these conditions the partial R^2s $\mathrm{IS}_i = \beta_i \gamma_i$ of course don't make any sense. We don't advocate the use of information shares if it is believed that Ω can be so far from diagonality. A credible identification is required, as otherwise observationally equivalent models might produce radically opposing results.

[6]According to (43) the off-diagonal element of Ω will be zero of $w = 0$ or $w = 1$, but the latter is outside the admissible range in (44).

[7]The exact formula is $\mathrm{IS}_1 = (4 - 8w + 3w^2)/(4 - 5w)$.

4 HIGHER-ORDER MODELS

In practice, microstructure data show second-order, and sometimes even higher-order, serial covariances. A natural way to model higher-order dynamics is by adding lagged noise terms e_{t-j} to the deviations from the efficient price.[8] Looking at the simplest case, the specification for the dealer behavior becomes

$$u_t = \alpha r_t + e_t + \Psi e_{t-1}, \tag{45}$$

with Ψ an $(N \times N)$ matrix. The moment conditions become

$$\begin{aligned} \Gamma_0 &= \mathrm{E}[\Delta p_t \Delta p_t'] = \sigma^2((\iota+\alpha)(\iota+\alpha)' + \alpha\alpha') + \Omega + (\Psi - I)\Omega(\Psi - I)' + \Psi\Omega\Psi', \\ \Gamma_1 &= \mathrm{E}[\Delta p_t \Delta p_{t-1}'] = -\sigma^2\alpha(\iota+\alpha)' + (\Psi - I)\Omega - \Psi\Omega(\Psi - I)', \\ \Gamma_2 &= \mathrm{E}[\Delta p_t \Delta p_{t-2}'] = -\Psi\Omega. \end{aligned} \tag{46}$$

As this is the model we will use in the empirical part of the paper, we analyze this specific case in a bit more detail. The additional parameter matrix Ψ is just identified from the second-order autocovariance matrix Γ_2. The random-walk variance is still over-identified as in the first-order case from the long-run covariance matrix $\sum_{j=-2}^{2} \Gamma_j = \sigma^2 \iota\iota'$. Identification of α is slightly more complicated than in the first-order model. Consider the following combination of moments:

$$D(\Gamma) = \Gamma_1' - \Gamma_1 + 2(\Gamma_2' - \Gamma_2) = \sigma^2(\alpha\iota' - \iota\alpha'). \tag{47}$$

This identifies α up to a translation along the unit vector. Like in the first-order case, the full set of equivalent solutions for α can be characterized by

$$\alpha = \tilde{\alpha} - w\iota, \tag{48}$$

where w is an arbitrary scalar and $\tilde{\alpha}$ is an initial admissible solution. As before, not all values for w are allowed, however, since the implied value for Ω has to be positive semidefinite. Given the other parameters, the noise covariance matrix Ω can be obtained from the moment equations; for example, using

$$\begin{aligned} \Gamma_1' + \Gamma_1 + 2(\Gamma_2' + \Gamma_2) &= -\sigma^2(\alpha\iota' + \iota\alpha' + 2\alpha\alpha') - \Omega - \Psi\Omega\Psi' \\ &= -\sigma^2(\alpha\iota' + \iota\alpha' + 2\alpha\alpha') - \Omega - \Gamma_2\Omega^{-1}\Gamma_2'. \end{aligned} \tag{49}$$

Unlike the first-order case (14), these moment equations are nonlinear in Ω due to the presence of Ω^{-1}. The identification rules for α of section 1 can also be applied in this case. The Watson restriction ($\pi'\alpha = 0$) and a diagonal Ω will lead to full identification. The definition of the information share in Equation (31) can then be applied directly.

[8] An alternative way to model higher-order dynamics is by including lagged effects of the efficient price in the transitory term. This imposes a particular structure on the serial correlation pattern, which may be at odds with the data. We therefore do not pursue this idea further.

The relation with the reduced form in the higher-order model is more complicated than in the first-order model. The reduced form of the second-order model can be written as a VMA(2) model,

$$\Delta p_t = \epsilon_t + B_1\epsilon_{t-1} + B_2\epsilon_{t-2}, \tag{50}$$

with $\mathrm{Var}(\epsilon) = \Sigma$. Cointegration requires that the coefficient matrices add up to

$$C(1) \equiv I + B_1 + B_2 = \iota\theta'. \tag{51}$$

Working out the moments gives:

$$\begin{aligned} \Gamma_0 &= \mathrm{E}[\Delta p_t \Delta p_t'] = \Sigma + B_1\Sigma B_1' + B_2\Sigma B_2', \\ \Gamma_1 &= \mathrm{E}[\Delta p_t \Delta p_{t-1}'] = B_1\Sigma + B_2\Sigma B_2', \\ \Gamma_2 &= \mathrm{E}[\Delta p_t \Delta p_{t-2}'] = B_2\Sigma. \end{aligned} \tag{52}$$

Substituting these values in the expression for $D(\Gamma)$ and using the cointegration restriction (51) gives

$$D(\Gamma) = (I - B_2)\Sigma\theta\iota' - \iota\theta'\Sigma(I - B_2) = \sigma^2(\alpha\iota' - \iota\alpha'). \tag{53}$$

From this equality, the full set of admissible values for α can be written as

$$\alpha = (I - B_2)\Sigma\theta - (w+1)\iota. \tag{54}$$

The expression for Ω is complicated, however, due to the presence of Ω^{-1} in (49). Notice that in the first-order case ($B_2 = 0$), the value $w = 0$ corresponds to the BN value ($\beta = \iota + \alpha = \Sigma\theta$).

Adding further lags $\Psi_j e_{t-j}$ does not alter anything in the identification of α. With more lags, the model becomes increasingly more difficult to analyze, but α remains easily connected to the asymmetry of the autocovariance structure. The result is given in the form of a theorem (see the Appendix for proof).

Theorem 2. *Let prices be generated by the unobserved components model (9) but with dealer shocks,*

$$u_t = \alpha r_t + \sum_{j=0}^{M} \Psi_j e_{t-j}, \tag{55}$$

where $\mathrm{E}[e_t r_s] = 0$ *for all* t *and* s. *Then*

$$\sum_{j=-(M+1)}^{M+1} \Gamma_j = \sigma^2\iota\iota' \tag{56}$$

and

$$\sum_{j=1}^{M+1} j(\Gamma_j' - \Gamma_j) = \sigma^2(\alpha\iota' - \iota\alpha'). \tag{57}$$

From these moment equations, the parameters σ^2 and the set of admissible values for α are easily found.

5 EMPIRICAL APPLICATIONS

We apply the unobserved components model to a set of Nasdaq dealer quotes, and compare with the VECM results. We discuss two applications. For the first, we consider midquotes of the five most active dealers in Intel for the six month period February–July 1999. With these data we consider the importance of ECNs relative to individual dealers in price discovery on Nasdaq. The second application uses midquote data for Expedia originating from Nasdaq and NYSE (ARCA) for the second half of 2007. For this data the basic question is about the information in quotes originating outside the Nasdaq system.

The purpose of the applications is to compare the alternative specifications. Does a UC model violate typical moments in high-frequency quote data? How stable are price discovery measures given the limited identification of the structural parameters? How reasonable is the over-identifying assumption of a diagonal matrix for individual deviations from the efficient price?

5.1 Intel 1999

The sample contains quotes that are sampled at two-minute intervals for 123 consecutive trading days in 1999. Since Intel is a liquid stock, there are hardly any missing values at this sampling frequency.[9] The five top dealers are the two ECN's Island (ISLD) and Instinet (INCA) and the three wholesale dealers Spear, Leeds & Kellogg Capital (SLKC); Mayer and Schweitzer (MASH); and Knight/Trimark Securities (NITE). The total number of observations for all series is 24,108.

Sample covariances are estimated omitting the overnight returns. The contemporaneous covariance matrix and the first two lags are reported in Table 2. Contemporaneous correlations among the quotes changes is only around 0.4. Since cointegration implies that the long-run correlation must be equal to one, enough dynamic structure remains despite the relatively low two minutes sampling frequency. All first-order autocorrelations are firmly negative. Second-order covariances are negligible, except for SLKC and NITE.

The variance of the random-walk component can be estimated from the long-run covariance matrix:

$$\bar{\Gamma} = \Gamma_0 + \sum_{i=1}^{L} (\Gamma_i + \Gamma_i') = \sigma^2 \iota\iota'. \tag{58}$$

[9] At higher frequencies we do not observe quote updates for the less active dealers in many time periods. Various ways to deal with these missings have been suggested; see, for example, Harris et al. (1995) and DeJong, Mahieu, and Schotman (1998). For clarity in this empirical illustration of the parameterization issues, we decided to keep the econometrics as simple as possible and work with data at the two-minute frequency.

Table 2 Data (auto-)covariances.

	Dealer	ISLD	INCA	SLKC	MASH	NITE
Lag 0	ISLD	5.37	*0.54*	*0.38*	*0.40*	*0.28*
(Γ_0)	INCA	2.36	3.50	*0.50*	*0.48*	*0.37*
	SLKC	2.37	2.49	7.08	*0.36*	*0.29*
	MASH	2.51	2.41	2.59	7.24	*0.27*
	NITE	1.84	1.95	2.21	2.07	7.98
Lag 1	ISLD	−1.31	0.08	0.02	0.02	0.01
(Γ_1)	INCA	0.22	−0.36	0.12	0.13	0.11
	SLKC	0.64	0.48	−1.21	0.42	0.29
	MASH	0.09	0.07	0.04	−2.06	−0.13
	NITE	0.52	0.41	0.31	0.44	−1.75
Lag 2	ISLD	−0.03	0.05	0.08	−0.01	0.01
(Γ_2)	INCA	0.02	−0.04	−0.03	0.02	−0.05
	SLKC	−0.11	−0.11	−0.44	−0.14	−0.14
	MASH	0.04	0.02	0.02	−0.09	0.12
	NITE	0.00	0.02	−0.02	−0.02	−0.58
Long run	ISLD	2.70	*1.01*	*0.94*	*0.94*	*0.80*
($\bar{\Gamma}$)	INCA	2.73	2.71	*0.94*	*0.94*	*0.81*
	SLKC	2.99	2.95	3.79	*0.87*	*0.75*
	MASH	2.64	2.65	2.90	2.93	*0.79*
	NITE	2.39	2.44	2.65	2.48	3.32
Information	ISLD	0.00	0.08	0.24	0.17	0.49
asymmetry	INCA	−0.08	0.00	0.21	−0.03	0.43
$D(\Gamma)$	SLKC	−0.24	−0.21	0.00	−0.14	0.27
	MASH	−0.17	0.03	0.14	0.00	0.28
	NITE	−0.49	−0.43	−0.27	−0.28	0.00
σ^2	2.54					

The table reports the sample covariances (*correlations*) for the time series of quote changes of the five most active dealers in Intel in the period February–July 1999. The entry on row i and column j for Γ_ℓ refers to the covariance $\mathrm{E}[\Delta p_{it}\Delta p_{j,t-\ell}]$. The long-run covariance matrix is defined as $\bar{\Gamma} = \Gamma_0 + \sum_{i=1}^{2}(\Gamma_i + \Gamma_i')$. The dealer information matrix is defined as $D(\Gamma) = \sum_{i=1}^{2} i(\Gamma_i' - \Gamma_i)/\sigma^2$. The scaling factor σ^2 is a GMM estimate from $\bar{\Gamma}$. Dealer acronyms are ISLD (Island), INCA (Instinet), SLKC (Spear, Leeds & Kellogg Capital), MASH (Mayer and Schweitzer), and NITE (Knight/Trimark Securities).

It is clear from Table 2 that with $L = 2$ not all elements in $\bar{\Gamma}$ are the same, nor that all correlations are equal to one. For the three wholesale dealers, and especially SLKC, the diagonal elements are still larger than for the two ECNs. Given the large number of observations, the differences are significant. Further lags must add some negative autocorrelations for the three dealers. We did not obtain full equality of all elements of $\bar{\Gamma}$ by adding a small number of lags. On the other hand, adding a few more lags hardly affects the estimate of the random walk variance σ^2. We therefore estimate all models with a maximum of second-order lags, with cointegration as a

Table 3 Vector error correction.

		Residual covariances (*correlations*)					Info shares	
Dealer	θ	ISLD	INCA	SLKC	MASH	NITE	min	max
ISLD	0.21	4.38	*0.70*	*0.52*	*0.55*	*0.37*	0.03	0.70
INCA	0.53	2.66	3.28	*0.61*	*0.61*	*0.44*	0.12	0.91
SLKC	0.10	2.64	2.66	5.82	*0.47*	*0.31*	0.01	0.52
MASH	0.09	2.71	2.60	2.67	5.55	*0.35*	0.01	0.51
NITE	0.04	2.01	2.04	2.12	2.15	6.61	0.01	0.26
		$\sigma^2 = 2.80$						

The table reports results obtained from the vector error correction model

$$\Delta p_t = c + As_{t-1} + D\Delta p_{t-1} + \epsilon_t$$

with $\mathrm{E}[\epsilon_t \epsilon_t'] = \Sigma$. The vector s_t contains the difference between the quotes of ISLD and each of the other four dealers. Parameters are estimated by OLS. The table reports estimates of the long-run impact matrix of the VECM,

$$C(1) = \iota\theta'.$$

The "Info shares" are the minimum and maximum information shares for each of the dealers, estimated using the methodology of Hasbrouck (1995). Residual correlations are in *italics*. The last entry in the table is the variance of the random-walk component, $\sigma^2 = \theta'\Sigma\theta$.

maintained hypothesis. Applying GMM to estimate σ^2 from the ten moments in $\bar{\Gamma}$ gives $\hat{\sigma}^2 = 2.54$ with a standard error of 0.06.

Implications for α can be obtained from the moment matrix

$$D(\Gamma) = \Gamma_1' - \Gamma_1 + 2(\Gamma_2' - \Gamma_2) = \sigma^2(\alpha\iota' - \iota\alpha'). \tag{59}$$

Elements of $D(\Gamma)$ scaled by σ^2 are reported in the last panel of Table 2. The sample moments in the table closely resemble the structure implied by (59). It appears that ISLD has the highest α_i, whereas SLKC and NITE have relatively small α_i.

For the formal analysis we first consider a reduced-form VECM, as this is the most commonly estimated model for inference on price discovery. We estimated the model with second-order dynamics,

$$\Delta p_t = c + As_{t-1} + D\Delta p_{t-1} + \epsilon_t, \tag{60}$$

where s_t is the vector of differences between the midquote of ISLD and each of the other four dealers, A a (5×4) matrix of error correction parameters, and D a (5×5) matrix. The most salient features of the VECM are reported in Table 3. The estimates of the information shares confirm the results of Huang (2002) that the ECNs dominate the price discovery on Nasdaq. Individual information shares of either ECNs or regular dealers are, however, in extremely wide intervals. For example, the lower and upper bound for ISLD are 3% and 70%, respectively.

Table 4 Vector-moving average.

Dealer	θ	Residual covariances (*correlations*)					Info shares	
		ISLD	INCA	SLKC	MASH	NITE	min	max
ISLD	0.25	4.31	*0.71*	*0.55*	*0.59*	*0.46*	0.07	0.75
INCA	0.49	2.66	3.23	*0.63*	*0.64*	*0.53*	0.14	0.90
SLKC	0.02	2.57	2.56	5.04	*0.51*	*0.44*	0.05	0.47
MASH	0.10	2.79	2.66	2.64	5.29	*0.42*	0.05	0.56
NITE	0.08	2.25	2.25	2.31	2.25	5.54	0.05	0.39
	$\sigma^2 = 2.64$				$J(20) = 103.21$			

The table reports results obtained from the vector-moving average model

$$\Delta p_t = B_2\epsilon_{t-2} + B_1\epsilon_{t-1} + \epsilon_t$$

with $\mathrm{E}[\epsilon_t\epsilon_t'] = \Sigma$ and under the cointegration restriction

$$C(1) = I + B_1 + B_2 = \iota\theta'.$$

Parameters are estimated by GMM using the moment conditions for Γ_0, Γ_1, and Γ_2. The "Info shares" are the minimum and maximum information shares for each of the dealers. Residual correlations are in *italics*. The last part of the table shows the variance of the random-walk component, $\sigma^2 = \theta'\Sigma\theta$, and the criterion value of the GMM estimator known as Hansen's J-statistic.

The wide intervals are caused by the strong contemporaneous correlations of the errors.[10]

The reduced form that is more closely related to the unobserved components model is the VMA. With second-order dynamics we estimate the model:

$$\Delta p_t = c + \epsilon_t + (\iota\theta' - I - B)\epsilon_{t-1} + B\epsilon_{t-2}. \tag{61}$$

The 45 parameters in θ, Σ, and B are estimated by GMM using the 65 moment conditions for Γ_0, Γ_1, and Γ_2. Table 4 shows estimation results.[11] Hansen's J-statistic rejects the 20 over-identifying moment conditions that result from the cointegration restriction $C(1) = \iota\theta'$. The empirical violation of this restriction in the model with second-order lags was already evident in Table 2. Although the VECM and VMA are not nested, it seems that the VMA fits the data better: all diagonal elements of Σ and also the determinant are smaller for the VMA.

Implications for the information shares are similar to the VECM results. All of the minimum, maximum, and θ are close to the VECM estimates. The high information share of INCA is mainly caused by its low residual variance.

[10] The wide intervals for the information shares are not an artifact of the sampling frequency: Huang (2002) finds similar wide intervals for Intel at the one-minute frequency. Huang (2002) uses slightly different data though, since he aggregates individual dealers into categories.

[11] The VMA representation in the table uses the invertible solution for the moment equations with all characteristic roots inside the unit circle except for the four unit roots imposed by cointegration.

Table 5 Unobserved components model.

Dealer	α	Error covariances Ω (*correlations*)					IS
		ISLD	INCA	SLKC	MASH	NITE	
		A. "Watson" representation: $\sum \alpha_i = 0$					
ISLD	0.074	1.160	*−0.47*	*−0.06*	*−0.13*	*−0.16*	0.262
INCA	0.027	−0.320	0.400	*−0.05*	*−0.20*	*−0.17*	0.490
SLKC	−0.008	−0.086	−0.049	2.088	*0.05*	*0.03*	0.049
MASH	0.024	−0.221	−0.191	0.115	2.401	*−0.07*	0.099
NITE	−0.116	−0.318	−0.194	0.073	−0.208	3.431	0.065
		$R^2 = 0.965$		$\sigma^2 = 2.64$		$J(20) = 103.21$	
		B. Approximately diagonal Ω					
ISLD	0.003	1.363	*−0.14*	*−0.04*	*−0.01*	*−0.07*	0.251
INCA	−0.044	−0.130	0.589	*0.14*	*−0.01*	*−0.02*	0.461
SLKC	−0.079	0.074	0.185	2.782	*0.10*	*0.09*	0.030
MASH	−0.047	−0.027	−0.011	0.260	2.587	*−0.02*	0.096
NITE	−0.187	−0.155	−0.030	0.281	−0.055	3.568	0.061
		$R^2 = 0.899$		$\sigma^2 = 2.64$		$J(20) = 103.21$	
		C. Diagonal covariance matrix					
ISLD	0.000	1.517					0.187
INCA	−0.008		0.626				0.446
SLKC	0.087			2.166			0.155
MASH	−0.032				2.489		0.123
NITE	−0.144					3.591	0.058
		$R^2 = 0.967$		$\sigma^2 = 2.54$		$J(29) = 162.03$	

The table reports results for the unobserved components model:

$$\begin{aligned} p_t &= \iota p_t^* + u_t, \\ p_t^* &= p_{t-1}^* + r_t, \\ u_t &= \alpha r_t + \Psi e_{t-1} + e_t. \end{aligned}$$

Panels A and B are reparameterizations of the VMA in Table 4. Panel A is the "Watson" representation with $\sum_i \alpha_i = 0$. Panel B reports the observationally equivalent representation with the lowest maximum correlation in $\Omega = \mathrm{E}[e_t e_t']$. In panel C diagonality of Ω is imposed. Entries report GMM estimates for σ^2, α, Ω, and the GMM criterion function. The *IS* column gives the information shares as defined in Equation (31). R^2 is the sum of the individual information shares, and equals the fraction of variance of the efficient price innovation explained by the observed prices.

By reparameterizing the VMA we obtain alternative observationally equivalent unobserved components representations with second-order dynamics as in (45). In Table 5 we report results for two of these equivalent models. The first is a model in "Watson" form ($\sum_i \alpha_i = 0$). In the second model, we have set w so that the maximum absolute correlation between the dealer noise terms e_{it} is minimal.

The latter model is the representation of the UC for which the noise covariance matrix is closest to diagonality. Since the models are observationally equivalent to the reduced-form VMA, they have the same GMM J-statistic.

The information shares from these structural models are within the minimum–maximum range of the reduced-form models. The two ECNs dominate with the information share of INCA being almost double that of ISLD. The dealer that contributes least to the price discovery process is NITE. It is the only dealer with a significantly different α_i.

In addition, Table 5 presents results for the over-identified model where diagonality of Ω has been imposed. Diagonality appears to be a good modeling assumption. Considering the large sample size, the restriction is only marginally rejected against the VMA (and its equivalent UC representations). It seems surprising that diagonality provides such a good fit to the data, since the correlation between the shocks of ISLD and INCA was -0.47 in the "Watson" model. Note, however, that shifting α in the direction of ι induces a compensating change in the structure of Ω. The results in panel B show that we can shift α such that Ω becomes almost diagonal with the maximum absolute correlation only 0.14.

The results for the "diagonal" model differ from the others mostly with regard to SLKC. In the diagonal model it has the highest α of all dealers. That is a somewhat surprising result, since from the matrix $D(\Gamma)$ in Table 2 we have seen that the raw covariances implied a low α relative to ISLD. The explanation is that the GMM weighting function also puts weight on fitting the total variance in Γ_0, for which it needs a much higher value of α. In the parsimonious diagonal model there are not enough other parameters to ease the tension between fitting the asymmetry in the lagged covariances between SLKC and other dealers and fitting the variance of SLKC quote updates.

Despite various possibilities for a more detailed modeling of these quote series, the main results seem robust across specifications. INCA is the most informative source for price discovery, followed by the other network Island (ISLD). The unobserved components model provides plausible point estimates of the information shares that are more informative than the VECM upper and lower bounds.

5.2 Expedia 2007

Data for our second application consist of midquotes for Expedia sampled at the one-minute frequency for the period July–December 2007. Quotes are taken from the TAQ database and split in three series depending on the origin of the quote. We distinguish between quotes from the Nasdaq and the NASD (TAQ codes Q, T, or D), the NYSE and NYSE-ARCA (TAQ codes N and P) and all other origins, which are mainly the NSX (TAQ code C) and the CBOE (TAQ codes I and W) exchanges. The overall fraction of quotes issued by these three groups is around 45% for NASD, 20% for NYSE and the remaining 35% for the other markets. It

Table 6 Data (auto-)covariances expedia 2007.

	Dealer	NYSE	NASD	REST
Lag 0	NYSE	0.938	*0.82*	*0.90*
(Γ_0)	NASD	0.888	1.261	*0.77*
	REST	0.890	0.884	1.035
Lag 1	NYSE	0.012	0.031	0.038
(Γ_1)	NASD	0.042	−0.149	0.046
	REST	0.041	0.039	−0.026
Lag 2	NYSE	0.006	0.007	0.013
(Γ_2)	NASD	0.015	0.013	0.015
	REST	0.003	0.002	−0.001
Long run	NYSE	0.975	*1.00*	*1.00*
($\bar{\Gamma}$)	NASD	0.984	0.987	*1.00*
	REST	0.985	0.985	0.981
Information	NYSE	0	0.027	−0.016
asymmetry	NASD	−0.027	0	−0.032
$D(\Gamma)$	REST	0.016	0.032	0
σ^2	0.984			

The table reports the sample covariances (*correlations*) for the time series of quote changes of the three alternative quote originations for Expedia in the period July–December 2007. The entry in row i, column j for Γ_ℓ refers to the covariance $E[\Delta p_{it} \Delta p_{j,t-\ell}]$. The long-run covariance matrix is defined as $\bar{\Gamma} = \Gamma_0 + \sum_{i=1}^{2}(\Gamma_i + \Gamma_i')$. The dealer information matrix is defined as $D(\Gamma) = \sum_{i=1}^{2} i(\Gamma_i' - \Gamma_i)/\sigma^2$. The long-run variance of the efficient price is a GMM estimate from $\bar{\Gamma}$. Market acronyms are NYSE (ARCA, NYSE), NASD (Nasdaq), and REST (all other).

is likely that most of the quotes in the third group originate from ECNs, who in recent years often report their trades through one of the regional exchanges.[12]

Tables 6–9 have the same structure as the previous Tables 2–5. From the data moments in Table 6, we see that contemporaneous correlations are much higher than for the data on individual dealers. This is so despite the higher sampling frequency (one-minute) than in the previous multiple dealer example. The cointegration restrictions in the long-run covariance matrix already hold very closely after adding two lags. Most interesting are the asymmetry measures, which indicate that NASD has the lowest α parameter and is therefore likely to be the least informative originator of quotes.

Tables 7 and 8 contain the estimates of the reduced-form VECM and VMA models. Consistent with the low value of α_{NASD} in the data moments, both models show very low estimates of θ_{NASD} indicating that shocks to NASD quotes do not contribute much to the innovation in the efficient price. Since contemporaneous

[12]See Goldstein et al. (2008), who study the importance of different trading venues for price discovery of Nasdaq-listed stocks. Their study looks at a variety of measures for 100 different stocks in the second quarter of 2003.

Table 7 Vector error correction expedia 2007.

		Residual cov (*corr*)			Info shares	
Dealer	θ	NYSE	NASD	REST	min	max
NYSE	0.51	1.00	*0.90*	*0.94*	0.024	0.974
NASD	0.05	0.96	1.15	*0.87*	0.001	0.824
REST	0.44	0.96	0.95	1.05	0.024	0.964
		$\sigma^2 = 1.00$				

The table reports results obtained from the vector error correction model

$$\Delta p_t = c + As_{t-1} + D\Delta p_{t-1} + \epsilon_t$$

with $\mathrm{E}[\epsilon_t \epsilon_t'] = \Sigma$. The vector s_t contains the difference between the quotes of NYSE and both other dealers. Parameters are estimated by OLS. The table reports estimates of the long-run impact matrix of the VECM,

$$C(1) = \iota\theta'.$$

The "Info shares" are the minimum and maximum information shares (percentage) for each of the dealers, estimated using the methodology of Hasbrouck (1995). Residual correlations are in *italics*. The last entry in the table is the variance of the random-walk component, $\sigma^2 = \theta'\Sigma\theta$.

Table 8 Vector-moving average expedia 2007.

		Residual cov (*corr*)			Info shares	
Dealer	θ	NYSE	NASD	REST	min	max
NYSE	0.69	0.93	*0.90*	*0.94*	0.04	0.99
NASD	0.05	0.90	1.08	*0.87*	0.00	0.82
REST	0.30	0.90	0.90	0.99	0.01	0.95
		$\sigma^2 = 0.966$		$J(20) = 6.91$		

The table reports results obtained from the vector moving average model

$$\Delta p_t = B_2\epsilon_{t-2} + B_1\epsilon_{t-1} + \epsilon_t$$

with $\mathrm{E}[\epsilon_t \epsilon_t'] = \Sigma$ and under the cointegration restriction

$$C(1) = I + B_1 + B_2 = \iota\theta'.$$

Parameters are estimated by GMM using the moment conditions for Γ_0, Γ_1, and Γ_2. The "Info shares"are the minimum and maximum information shares for each of the dealers. Residual correlations are in *italics*. The last part of the table shows the variance of the random-walk component, $\sigma^2 = \theta'\Sigma\theta$, and the criterion value of the GMM estimator known as Hansen's J-statistic.

correlations among the shocks are high, alternative Cholesky factorizations lead to very wide bands for the price discovery measures. In all cases the estimates are completely uninformative with lower bounds close to zero and upper bound approximately one.

Table 9 Unobserved components model.

Dealer	α	NYSE	NASD	REST	IS
		Noise cov/*corr* Ω			
		A. "Maximum" representation max $\sum \alpha_i$			
NYSE	−0.055	0.011	*−0.44*	*−0.51*	0.685
NASD	−0.077	−0.019	0.156	*−0.10*	0.110
REST	−0.044	−0.014	−0.010	0.067	0.205
	$R^2 = 0.997$	$\sigma^2 = 0.98$		$J(6) = 6.913$	
		B. Approximately diagonal Ω			
NYSE	−0.060	0.027	*0.09*	*−0.09*	0.670
NASD	−0.082	0.006	0.199	*0.04*	0.110
REST	−0.049	−0.004	0.004	0.074	0.205
	$R^2 = 0.980$	$\sigma^2 = 0.98$		$J(6) = 6.913$	
		C. Diagonal covariance matrix			
NYSE	−0.106	0.027			0.670
NASD	−0.126		0.186		0.095
REST	0.098			0.080	0.235
	$R^2 = 0.978$	$\sigma^2 = 1.00$		$J(8) = 7.887$	

The table reports results for the unobserved components model:

$$\begin{aligned} p_t &= \iota p_t^* + u_t, \\ p_t^* &= p_{t-1}^* + r_t, \\ u_t &= \alpha r_t + \Psi e_{t-1} + e_t. \end{aligned}$$

Panels A and B are reparameterizations of the VMA in Table 8. Panel A is the representation for which $\sum_i \alpha_i$ is at the maximum admissible value. In panel B, diagonality of Ω is imposed. Entries report GMM estimates for σ^2, α, Ω, and the GMM criterion function. The IS column gives the information shares as defined in Equation (31). R^2 is the sum of the individual information shares, and equals the fraction of variance of the efficient price innovation explained by the observed prices.

The reduced-form VMA model is equivalent to a range of specifications for the unobserved components model with different values for $\sum_i \alpha_i$. The Watson representation is not in this equivalence set. The equivalent UC model with the largest possible $\sum_i \alpha_i$ has negative α for all quote origins and $\sum_i \alpha_i = -0.17$. Even with the large number of 45,356 observations, the restriction of a diagonal Ω cannot be rejected. Information shares for all equivalent UC models are plotted in Figure 1. From the figure we conclude that all representations, which differ by their identifying restriction on $\sum_i \alpha_i$ shown on the horizontal axis, lead to almost identical estimates of the information shares. What is analytically a problem of under-identification, is empirically a negligible effect. There is only a small range of admissible values for $\sum_i \alpha_i$, and the different values do not affect the economically

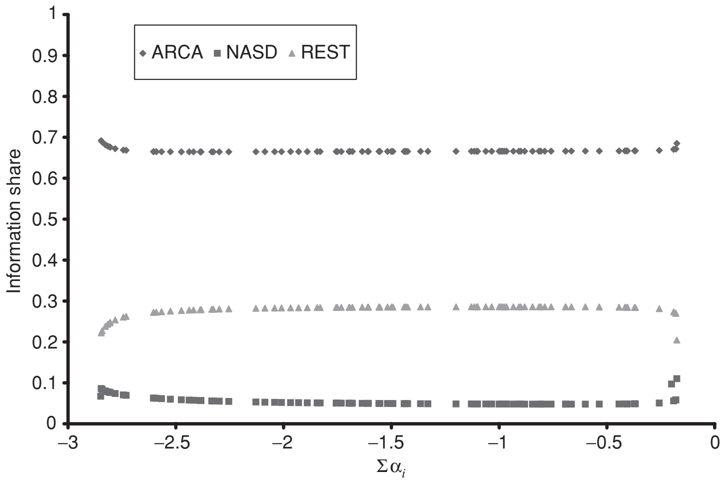

Figure 1 Identification and information shares. The figure shows the information shares for alternative, observationally equivalent, estimates of the structural parameters of the unobserved components model for the EXPEDIA data. Each point in the figure has been obtained by estimating the model for a different value of $\sum \alpha_i$, shown on the horizontal axis. All estimates achieve the same value of the GMM criterion.

interesting parameters related to the informativeness of the different quote series for the price discovery process.

6 CONCLUSION

In this paper we proposed an Unobserved Components model for price discovery in fragmented markets. The model decomposes the observed prices in an underlying common efficient price and market-specific transitory components. We show how this model is related to the usual VAR or VECM models for price discovery, and argue that the unobserved components model is a natural and parsimonious way of modeling price discovery. The parameters in the unobserved components model have natural interpretations as the variance of the efficient price, variances, and covariances of the transitory terms, and correlations between transitory terms and the efficient price. Because of this structure, it is easy to impose economically interesting or plausible restrictions on the model; for example, diagonality of the transitory noise covariance matrix. Moreover, the dynamic structure (lag length) of the model can be easily adapted to the serial correlation pattern observed in the data.

We also propose a new measure for the contribution to price discovery based on a permanent/transitory decomposition of the error terms instead of the usual Cholesky decomposition. This measure is based on the covariance between the transitory components and the efficient price.

Our empirical results using Nasdaq quotes show that the approach leads to meaningful and informative estimates of information shares. We conclude that the key parameters of interest can be estimated from a parsimonious unobserved components model. These parsimonious models could prove useful for applications on smaller data sets; for example, around specific events such as corporate announcements.

APPENDIX: PROOFS

BN Representation Has Maximum α

We prove the assertion on section 1.2 that the BN normalization of α is the highest possible value in the random walk plus noise UC model. Substituting the BN expressions (21) and (22) in the solution set for Ω, we find

$$\Omega = \Sigma - \Sigma\theta\theta'\Sigma/\sigma^2 + w(\iota\theta'\Sigma + \Sigma\theta\iota') - w(w+1)\sigma^2\iota\iota'. \tag{A1}$$

We now show that this implies that only positive values for w are allowed. First, pre- and post-multiply the expression for Ω by θ and use $\theta'\Sigma\theta = \sigma^2$ to obtain

$$\theta'\Omega\theta = 2w\sigma^2\Theta - w(w+1)\sigma^2\Theta^2, \tag{A2}$$

where $\Theta = \iota'\theta$ is the sum of elements of θ. The right-hand side of this equation is a quadratic function of w with roots $w_1 = 0$ and

$$w_2 = \frac{2}{\Theta} - 1. \tag{A3}$$

As long as $0 < \Theta < 2$, w_2 is positive and $\theta'\Omega\theta$ is positive for values $0 < w < w_2$. Negative values for w are not allowed, like too high positive values (too low values of α). The condition $0 < \Theta < 2$ seems plausible. Individual elements of θ will likely be positive if innovations to prices are positively correlated with an innovation in the efficient price. Furthermore, consider the time-series process for $q_t = \Theta^{-1}\theta' p_t$, a weighted average of the prices with positive weights,

$$\Delta q_t = \Theta^{-1}\theta'\epsilon_t - \Theta^{-1}\theta'(I - \iota\theta')\epsilon_{t-1}, \tag{A4}$$

which can be written as

$$\Delta q_t = e_t - (1-\Theta)e_{t-1}, \tag{A5}$$

with $e_t = \Theta^{-1}\theta'\epsilon_t$. An MA coefficient $1-\Theta$ between 0 and 1 seems reasonable for stationary microstructure data with negative first-order serial correlation. If $\Theta = 1$, then q_t is a weighted average of individual prices that follows a random walk, equal to the efficient price p_t^*. In the empirical applications we always find that $0 < \Theta < 1$, and usually Θ close to one.

Proof of Theorem 1

Use the matrix inversion lemma

$$\Upsilon^{-1} = \Omega^{-1} - \frac{\sigma^2}{1+\sigma^2(\beta'\Omega^{-1}\beta)}\Omega^{-1}\beta\beta'\Omega^{-1}$$

to compute

$$\gamma = \Upsilon^{-1}\beta\sigma^2 = \frac{\sigma^2}{1+\sigma^2(\beta'\Omega^{-1}\beta)}\Omega^{-1}\beta,$$

and thus

$$\beta_j\gamma_j = \frac{\sigma^2}{1+\sigma^2(\beta'\Omega^{-1}\beta)}\beta_j(\Omega^{-1}\beta)_j,$$

which can be rewritten in the form given in the theorem.

Proof of Theorem 2

The representation for the price change is

$$\Delta p_t = (\iota+\alpha)r_t - \alpha r_{t-1} + \Psi_0 e_t + \sum_{i=1}^{M}(\Psi_i - \Psi_{i-1})e_{t-i} - \Psi_M e_{t-M-1}. \tag{A6}$$

The identification of σ^2 in (56) is a general result, which follows directly from substituting the moment equations. For the second result, we start by analyzing the covariance structure of the series $\Psi_0 e_t + \sum_{j=1}^{M}(\Psi_j - \Psi_{j-1})e_{t-j} - \Psi_M e_{t-M-1}$. The autocovariances are

$$\begin{aligned}
\tilde{\Gamma}_{M+1} &= -\Psi_M\Psi_0',\\
\tilde{\Gamma}_M &= -\sum_{i=M-1}^{M}\Psi_i\Psi_{i-M+1}',\\
\tilde{\Gamma}_j &= (\Psi_j - \Psi_{j-1})\Psi_0' + \sum_{i=j+1}^{M}(\Psi_i - \Psi_{i-1})(\Psi_{i-j} - \Psi_{i-j-1})' - \Psi_M(\Psi_{M-j+1} - \Psi_{M-j})'\\
&= -\sum_{i=j-1}^{M}\Psi_i\Psi_{i-j+1}' + 2\sum_{i=j}^{M}\Psi_i\Psi_{i-j}' - \sum_{i=j+1}^{M}\Psi_i\Psi_{i-j-1}' \qquad 1 < j < M.
\end{aligned} \tag{A7}$$

Summing the elements in (A7) gives

$$\sum_{j=1}^{M+1} j\tilde{\Gamma}_j = -\sum_{i=0}^{M}\Psi_i\Psi_i', \tag{A8}$$

since all terms of the form $\sum_{i=j}^{M}\Psi_i\Psi_{i-j}'$ cancel because the coefficients $(-(j+1)+2j-(j-1))$ are always zero. Putting the efficient price changes $(\iota+\alpha)r_t - \alpha r_{t-1}$

back in, the same sum of the moments of Δp_t follows as

$$\sum_{j=1}^{M+1} j\Gamma_j = -\sigma^2\alpha(\iota + \alpha)' - \sum_{i=0}^{M} \Psi_i \Psi_i'. \tag{A9}$$

Subtracting the transpose of this matrix, all symmetric terms cancel and we are left with the result

$$\sum_{j=1}^{M+1} j(\Gamma_j' - \Gamma_j) = \sigma^2(\alpha\iota' - \iota\alpha'). \tag{A10}$$

Received September 4, 2007; revised March 27, 2008; accepted July 31, 2009.

REFERENCES

Bandi, F., and J. Russell. 2006. Separating Microstructure Noise from Volatility. *Journal of Financial Economics* 79: 655–692.

Covrig, V., and M. Melvin. 2002. Asymmetric Information and Price Discovery in the FX Market: Does Tokyo Know More about the Yen? *Journal of Empirical Finance* 9: 271–285.

De Jong, F., R. Mahieu, and P. Schotman. 1998. Price Discovery in the Foreign Exchange Market. *Journal of International Money and Finance* 17: 5–27.

Dittmar, R., and K. Yuan. 2008. Do Sovereign Bonds Benefit Corporate Bonds in Emerging Markets? *Review of Financial Studies* 21: 1983–2014.

Glosten, L., and L. Harris. 1988. Estimating the Components of the Bid/Ask Spread. *Journal of Financial Economics* 21: 123–142.

Goldstein, M., A. Shkilko, B.V. Ness, and R. V. Ness. 2008. Competition in the Market for NASDAQ Securities. *Journal of Futures Markets* 11: 113–143.

Harris, F., T. McInish, G. Shoesmith, and R. Wood. 1995. Cointegration, Error Correction, and Price Discovery on Informationally Linked Security Markets. *Journal of Financial and Quantitative Analysis* 30: 563–579.

Harris, F., T. McInish, and R. Wood. 2002. Security Price Adjustments across Exchanges: An Investigation of Common Factor Components for Dow Stocks. *Journal of Financial Markets* 5: 277–308.

Harvey, A. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Hasbrouck, J. 1993. Assessing the Quality of a Security Market: A New Approach to Transaction-Cost Measurement. *Review of Financial Studies* 6: 191–212.

Hasbrouck, J. 1995. One Security, Many Markets: Determining the Contributions to Price Discovery. *Journal of Finance* 50: 1175–1199.

Huang, R. 2002. The Quality of ECN and NASDAQ Market Maker Quotes. *Journal of Finance* 57: 1285–1319.

Hupperets, E., and A. Menkveld. 2002. Intraday Analysis of Market Integration: Dutch Blue Chips Traded in Amsterdam and New York. *Journal of Financial Markets* 5: 57–82.

Lehmann, B. 2002. Some Desiderata for the Measurement of Price Discovery across Markets. *Journal of Financial Markets* 5: 259–276.

Madhavan, A. 2000. Market Microstructure: A Survey. *Journal of Financial Markets* 3: 205–258.

Menkveld, A., S. Koopman, and A. Lucas. 2007. Modeling Around-the-Clock Price Discovery for Cross-Listed Stocks Using State Space Methods. *Journal of Business and Economics Statistics* 25: 213–225.

Mizrach, B., and C. Neely. 2008. Information Shares in the US Treasury Market. *Journal of Banking and Finance* 32: 1221–1233.

Morley, J., C. Nelson, and E. Zivot. 2003. Why Are the Beveridge–Nelson and Unobserved-Components Decompositions of GDP So Different? *Review of Economics and Statistics* 85: 235–243.

Sims, C. 1980. Macroeconomics and Reality. *Econometrica* 48: 1–48.

Taylor, N. 2008. "Time-Varying Price Discovery in Fragmented Markets." Working paper, Cardiff University.

Upper, C., and T. Werner. 2002. "Tail Wags Dog? Time-Varying Information Shares in the Bund Market." Discussion paper 24/02, Economic Research Centre of the Deutsche Bundesbank.

Zhou, B. 1996. High-Frequency Data and Volatility in Foreign-Exchange Rates. *Journal of Business and Economics Statistics* 14: 45–52.